

Beyond bias and discrimination: redefining the AI ethics principle of fairness in healthcare machine-learning algorithms

Year: 2022 | Citations: 113 | Authors: B. Giovanola, S. Tiribelli

Abstract

The increasing implementation of and reliance on machine-learning (ML) algorithms to perform tasks, deliver services and make decisions in health and healthcare have made the need for fairness in ML, and more specifically in healthcare ML algorithms (HMLA), a very important and urgent task. However, while the debate on fairness in the ethics of artificial intelligence (AI) and in HMLA has grown significantly over the last decade, the very concept of fairness as an ethical value has not yet been sufficiently explored. Our paper aims to fill this gap and address the AI ethics principle of fairness from a conceptual standpoint, drawing insights from accounts of fairness elaborated in moral philosophy and using them to conceptualise fairness as an ethical value and to redefine fairness in HMLA accordingly. To achieve our goal, following a first section aimed at clarifying the background, methodology and structure of the paper, in the second section, we provide an overview of the discussion of the AI ethics principle of fairness in HMLA and show that the concept of fairness underlying this debate is framed in purely distributive terms and overlaps with non-discrimination, which is defined in turn as the absence of biases. After showing that this framing is inadequate, in the third section, we pursue an ethical inquiry into the concept of fairness and argue that fairness ought to be conceived of as an ethical value. Following a clarification of the relationship between fairness and non-discrimination, we show that the two do not overlap and that fairness requires much more than just non-discrimination. Moreover, we highlight that fairness not only has a distributive but also a socio-relational dimension. Finally, we pinpoint the constitutive components of fairness. In doing so, we base our arguments on a renewed reflection on the concept of respect, which goes beyond the idea of equal respect to include respect for individual persons. In the fourth section, we analyse the implications of our conceptual redefinition of fairness as an ethical value in the discussion of fairness in HMLA. Here, we claim that fairness requires more than non-discrimination and the absence of biases as well as more than just distribution; it needs to ensure that HMLA respects persons both as persons and as particular individuals. Finally, in the fifth section, we sketch some broader implications and show how our inquiry can contribute to making HMLA and, more generally, AI promote the social good and a fairer society.